

The NIST Chemistry WebBook

The NIST Chemistry WebBook is a site on the Internet (<http://webbook.nist.gov/>) that distributes physical and chemical property data on well-defined chemical species and reactions. The purpose of this activity is to improve the nation's science and technology efforts by providing convenient access to NIST Standard Reference Data[†] to scientists, engineers, and the general public.

P.J. Linstrom and W.G. Mallard (Div. 838)

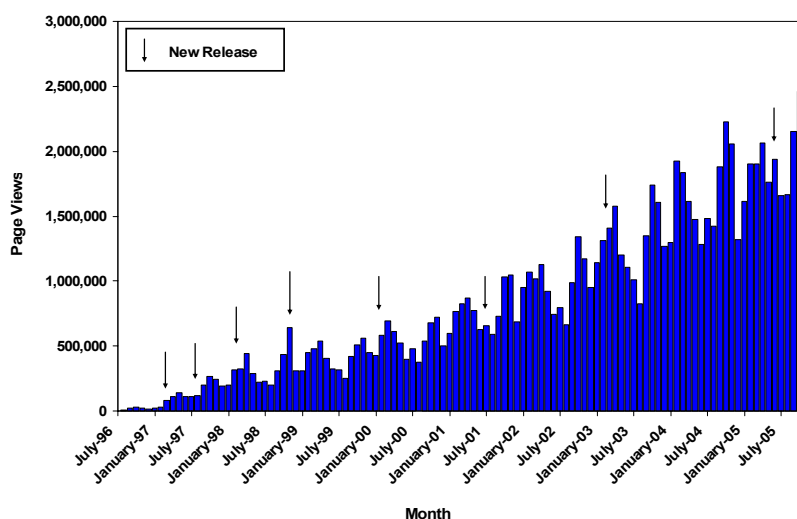
The data in the WebBook are from databases developed by NIST and outside contributors. Such data were traditionally published only in hardcopy format. With the widespread introduction of personal computers SRD products were distributed in the form of databases on floppy disks and CD-ROMs. The development of the Internet provided a third avenue for distribution of NIST data. The NIST Chemistry WebBook was one of the first efforts to distribute these data over the Internet. This has several advantages for NIST's customers: rapid access to data from any location with Internet access, updates and corrections made without any action required by the user, and better integration of data and relevant documentation. The NIST Chemistry WebBook is a service, that it is available on a continuing basis. Users worldwide can (and do) access the site 24 hours a day.

The greatest technical challenge in this effort is the integration of databases from multiple sources into a single collection of data. Because of the lack of standards for physical and chemical properties the data collections included in the site use diverse formats and conventions for reporting data. Problems with integrating these data sets are exacerbated by traditional chemical informatics problems associated with the identification of chemical species: inconsistent nomenclature, ambiguous or erroneous references to third party registries, and inconsistent conventions for chemical formulas. Systematically resolving these issues provides a common infrastructure for searching, identifying, and providing reliable information. Additionally, the WebBook employs internationally recognized developments in nomenclature, data exchange standards, and chemical identifiers, such as InChI.

The site was established in 1996 and has grown to encompass a wide variety of thermochemical, ion energetics, physical, solubility, spectroscopic, and chromatographic data. The site also includes two special features: a set of interactive physical property models developed at the NIST Boulder labs and a group-additivity-based estimator for gas-phase thermodynamic properties. The physical property models provide thermodynamic and transport property data at user-supplied conditions for a number of industrially important fluids. The group-additivity-based property starts with user-supplied structures using Benson increments. The NIST Chemistry WebBook is part of the NIST Standard Reference Data Program and is the result of the work of many contributors. These are identified in the "credits" section of the web page at <http://webbook.nist.gov/chemistry/>.

The NIST Chemistry WebBook is the most widely used SRD product developed by NIST. It is used by scientists and engineers in industry and academia and by students at the high school, college, and graduate school levels.

As noted in the figure, usage of the site has grown over time. The site now often receives more than two million page views per month.



The release of the June 2005 edition of the site was a major achievement for FY 2005. This release included the addition of the NIST retention index (gas chromatography)

[†] cf. United States Congress in the Standard Reference Data Act (United States Code, title 15, chapter 7A, section 290).

database (for approximately 27,000 compounds), and updates to existing data collections. The content of this release is found at

<http://webbook.nist.gov/chemistry/history.html>.

Finally, two major functional changes to the site were carried out this year. The first is the public release of translations developed by the EU-funded program "Eurospec - Access to Research Spectroscopic Data and Associated Chemical Knowledge - GTC1-2001-43000". These translations make the site accessible to speakers of Spanish, Portuguese, French, and Czech. In addition to translation work carried out in Europe, significant work was done at NIST that supports multiple-language publication on the site.

The second major improvement was to incorporate support for the IUPAC International Chemical Identifier (InChI) in the site. InChI is a new standard for identifying chemical species based on their structure. Addition of InChI has allowed easy identification (and linking) of isomers and isotopomers. The site now publishes InChI strings for every species with a chemical structure and allows searching in InChI strings as an alternate chemical name.

Future Plans:

This is an ongoing activity. Enhancements to the site's data collections and functionality occur continually over time. The following items are expected for FY 2006:

1. Improved handling and display of structures for repeating units in monomers.
2. Improved support for the IUPAC International Chemical Identifier, including the ability to display information derived from the identifier for species not in the database.
3. Expansion of retention index data to include polar-phase columns.

It is anticipated that components of the NIST Chemistry WebBook will be used in future chemical informatics efforts in CSTL. Plans are already underway to use the chemical structure display capabilities of the site in databases being developed to identify specific, structure-based reaction classes and use these classes in systematic development of reaction models for combustion.

The NIST Chemistry WebBook already indexes several of NIST's on-line databases including the gas-phase kinetics database, the computation chemistry comparison and benchmark database, the NIST atomic spectra database, the NIST microwave spectra database, the electron-impact ionization cross section database, and the thermophysical properties of hydrogen web site. A major planned application of the infrastructure of the NIST Chemistry WebBook is to expand this index to cover all data for chemical species, mixtures, and reactions produced by the Physical and Chemical Properties Division and other related databases. This index will support a "data portal" which will aid our customer by providing:

1. A single point of access for NIST's databases in this area;
2. A more complete awareness of currently less-well-known databases;
3. A user-friendly, unified data-access interface.

P.J. Linstrom and W.G. Mallard, Eds., **NIST Chemistry WebBook**, NIST Standard Reference Database Number 69, June 2005, National Institute of Standards and Technology, Gaithersburg MD, 20899 (<http://webbook.nist.gov>).